Introduction To Experimental Design

Advantages

- Gain scientific knowledge
- Experimenter can have full control
- Assign experimental units at random to different treatments
- Administer treatments and observe effects

Basic concepts

- **Response variable** – aim is to study effect of treatment on response variable – also called the **dependent variable**
- **Experimental unit** – an entity which can be treated separately and which can be assigned individually at random to treatment group
- **Treatment** – a set of actions to be performed on experimental units, called levels of treatment
- **Inherent variability** – an error term representing the random part of each observation
- **Blocking** – to divide experimental units into uniform groups (blocks)

More basic concepts

- Analysis of covariance – variables that are correlated with response variable
- Randomization – to avoid bias – deliberately distributing experimental units randomly into groups
- Independent variable. An independent variable (also called a factor or treatment) is an explanatory variable manipulated by the experimenter.
- Dependent variable. Say the researcher is looking at the effect of vitamins on health. The dependent variable in this experiment would be some measure of health (annual doctor bills, number of colds caught in a year, number of days hospitalized, etc.).
- Confounding occurs when the experimental controls do not allow the experimenter to reasonably eliminate plausible alternative explanations for an observed relationship between independent and dependent variables.

Example

Identify the **experimental units** or **subjects, treatments, factors, factor levels** and **response variable.**

In a Food technology study involving the storage of frozen strawberries, 10 pints were stored at each of 5 storage times. Storage times were randomly assigned to the pints. The amount of ascorbic acid content for each pint was measured after storage.

Characteristics of a Well-Designed Experiment

- Control. Control refers to steps taken to reduce the effects of extraneous variables (i.e., variables other than the independent variable and the dependent variable). These extraneous variables are called lurking variables.

  Control involves making the experiment as similar as possible for subjects in each treatment condition. Three control strategies are control groups, placebos, and blinding.
- Randomization. Randomization refers to the practice of using chance methods (random number tables, flipping a coin, etc.) to assign subjects to treatments. In this way, the potential effects of lurking variables are distributed at chance levels (hopefully roughly evenly) across treatment conditions.
- Replication. Replication refers to the practice of assigning each treatment to many experimental subjects. In general, the more subjects in each treatment condition, the lower the variability of the dependent measures.

Characteristics of a Well-Designed Experiment

- Placebo. Often, subjects respond differently after they receive a treatment, even if the treatment is neutral. A neutral treatment that has no "real" effect on the dependent variable is called a placebo, and a subject's positive response to a placebo is called the placebo effect.

  To control for the placebo effect, researchers often administer a neutral treatment (i.e., a placebo) to the control group. E.g. using a sugar pill in drug research.
- Blinding. Of course, if subjects in the control group know that they are receiving a placebo, the placebo effect will be reduced or eliminated; and the placebo will not serve its intended control purpose.

  Blinding is the practice of not telling subjects whether they are receiving a placebo. Often, knowledge of which groups receive placebos is also kept from people who administer or evaluate the experiment. This practice is called double blinding

Characteristics of a Well-Designed Experiment

- **Causation**. It allows the experimenter to make causal inferences about the relationship between independent variables and a dependent variable.
- **Variability**. It reduces variability within treatment conditions, which makes it easier to detect differences in treatment outcomes.

Assumptions of ANOVA

- Additivity – experimental designs follow a mathematical linear model:

  $X_i = \mu + t_i + e_i$ (completely randomized design)

  where
  X_i = the value of the experimental unit
  μ = the general mean
  t_i = the treatment effect
  e_i = error term

Assumptions of ANOVA

- Normality

  Errors must be normally distributed.
  Validity depends on the measure chosen.
- Support for the normality assumption:

  Central limit theorem (As n becomes sufficiently large – the sampling distribution of means ten toward a normal distribution)

  F-Test: (which is used by ANOVA) probabilities of errors (type I & II) are not affected "severely" by "moderate" departures from normality

  How to rectify
  Use transformation of measures

Assumptions of ANOVA

- Independent errors

  I.e. the error of one observation is not correlated with the error of another observation

Assumptions of ANOVA

- Homogeneous variance

  Experimental error variance should be constant over all observations (This is usually reduced if treatments have same number of replicates)

  How to rectify:
  i) Careful consistent administering of treatments
  ii) Type of measurement taken (assumption of normality)
  iii) Treatment levels

How to detect failures of assumptions

- Additivity – use Tukey's test
- Normality – calculate errors for each observation, check for normality
- Homogeneity of variance – calculate mean and variance for each treatment – plot on scatter diagram

Reducing error

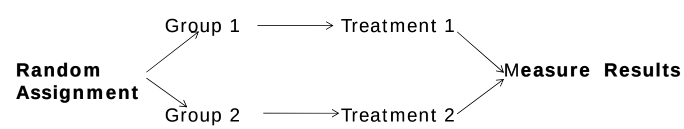
- Increase size of experiment – more replications
- Refining experimental conditions
- Reducing variability in experimental material

Standard experimental designs

- Completely randomized design
- Randomized complete block design
- Randomized blocks without replicates
- Incomplete randomized block design
- Latin square designs
- Graeco-Latin square design

Completely randomized design (CRD)

- An equal number of experimental units are assigned randomly to a number of treatment levels
  e.g. If there are 3 treatment levels and there are 90 units, then 30 units are randomly assigned to each treatment level
- Only when there is no blocking factor



First example of CRD

Blanching

A process development team conducted a study to establish the effect of blanching on the vitamin C content of mango slices.

Methodology

Four different methods (hot water, steam, microwave, autoclave) were used and applied for 5 minutes to *randomly* selected mango slices. The blanched mango slices were then sampled for Vitamin C content determination in triplicates

Experimental objective:

To establish if one blanch method is superior for vitamin C retention than the others.

First example of CRD : Blanching - Results

Blanching method			
Hot water	Steam	Microwave	Autoclave
77	75	74	69
73	70	70	74
78	74	72	67

Vitamin C content of mango slices (mg/100g)

Second example of CRD: Smoke Intensity

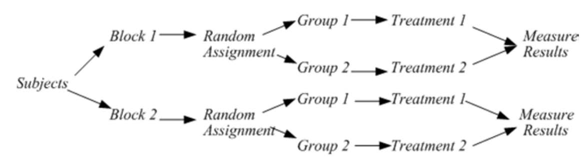
The table shows the smoke intensity data from a sensory test for three formulations of a food product obtained using a 7-point scale.

Experimental Objective:
The researcher wanted to find out whether there are significant differences among formulations with respect to the level of smoke using a 7-point scale where 1 denotes “zero amount” and 7 denotes “extremely strong” amount.

Formulation		
A	B	C
5	7	4
4	6	5
6	4	7
5	7	4
5	5	4
7	4	3
4	7	3
3	7	4
4	7	5

Randomized blocks

- Divide experimental units into uniform blocks
- Randomization performed within the blocks



Randomized blocks without replicates

- As many experimental units per block as treatment levels i.e. 3 treatment levels, and only three experimental units

Complete Randomized Block Design: Example – Smoke Intensity

If we decided to block the variation in scores among the panelists we would have a randomized block design. In which case the treatment and panel is partitioned, thus leaving a clear treatment effect measure. The assessor effect may or may not be tested for significance.

	Formulation		
Panelists	A	B	C
1	5	7	4
2	4	6	5
3	6	4	7
4	5	7	4
5	5	5	4
6	7	4	3
7	4	7	3
8	3	7	4
9	4	7	5

Complete Randomized Block Design: Example – Apple pie

Four different recipes (A, B, C, D) were tested for making apple pie. In order to evaluate the products a sensory testing panel requires four pies of each recipe.

The pilot plant is unable to bake more than four pies in a day. The most obvious solution would be to bake one pie of each type (A, B, C, and D) on one day, and submit them to the testing panel each day for four days.

The problem with this setup is that it does not take into account day-to-day variations that may occur among the judges or changes in the laboratory personnel making the pies.

These are extraneous variables. The conclusions from such a test might improperly disclose apparent differences between pies which are truly differences between days. Each pie has been baked four times over the four-day period.

Complete Randomized Block Design: Example – Apple pie

Day 1	Day 2	Day 3	Day 4
B	A	C	C
D	C	A	D
A	B	D	A
C	D	B	B

Each pie has been baked four times over the four-day period.

Another example

The article "Responsiveness of Food Sales to Shelf Space Changes in Supermarkets" (*Journal of Marketing Research [1964]: 63–67)* described an experiment to assess the effect of allotted shelf space on product sales. Two of the products studied were baking soda (a staple product) and Tang (considered to be an impulse product). Six stores (blocks) were used in the experiment, and six different shelf-space allotments were tried for 1 week each. Space allotments of 2, 4, 6, 8, 10, and 12 ft were used for baking soda, and 6, 9, 12, 15, 18, and 21 ft were used for Tang.

The author speculated that sales of staple goods would not be sensitive to changes in shelf space, whereas sales of impulse products would be affected by changes in shelf space.

http://academic.cengage.com/resource_uploads/downloads/0495390879_94372.pdf

Store	Baking Soda Shelf Space					
	2	4	6	8	10	12
1	36	42	36	40	30	22
2	74	61	65	67	83	84
3	40	58	42	73	69	63
4	43	65	65	41	43	47
5	27	33	35	17	40	26
6	23	31	36	38	42	37

Store	Tang Shelf Space					
	6	9	12	15	18	21
1	30	35	25	25	38	31
2	47	59	43	62	65	48
3	47	55	48	54	36	54
4	29	19	41	27	33	39
5	17	11	25	23	24	26
6	22	9	19	18	25	22

The Effect of Homogenization, Stabilizer and Amylase Treatment on Viscosity of Passion Fruit Juice.

A. R. Kahawa *B.Sc.,(Hons) M.Sc., University of Nairobi; **M. W. Okoth**** Ph.D., ETH Zurich; **J. K. Imungi**. Ph.D, Cornell University.

A Randomized Block design experiment with two blocks and three treatments was used. The two blocks were "before" and "after" pasteurization respectively. The three treatments were represented by homogenization, stabilizer and addition of amylase. Preliminary investigations were analyzed using Least Significant Test (LSD) to determine the minimum and maximum levels of each factor which give significant changes (P≤.100) in viscosity.

Vol. 4 No. 1 The Journal of Food Technology in Africa 22

Complete Randomized Block Design: Example – Bread Loaf Formulas

Four bread formulas are compared for loaf volume when baked in a microwave oven.

The goal is to compare formulas, not microwave ovens

	Oven		
Formula	GE	TA	SA
1	126	143	160
2	150	151	157
3	112	137	126
4	162	151	175

Incomplete Randomized Block Design:

Assume that we wish to compare 7 pie recipes.

The sensory panel judges can evaluate no more than 3 pies before losing their ability to detect flavor differences. The 7 (A through G) recipes for pies in the incomplete block are indicated in the table.

Note that each judge evaluates only 3 pies and that each pie is sampled by 3 different judges. The design is carefully balanced so that each group appears the same number of times and each formula appears the same number of times

Judges						
1	2	3	4	5	6	7
A	G	A	C	B	C	D
E	F	B	G	E	D	G
C	A	D	B	F	F	E

Factorial Experiments

- Where 2 or more factors have to be tested
- With 2 or more levels for each factor
- The response using various combinations of factors and levels are measured
- Researcher interested in effects of the individual treatments (Main effects)
- Researcher is also interested in the effects of the treatments combined (Interaction)

Advantages of factorial experiments

- Quicker and more economic to include several treatments in one experiment
- Level of precision and accuracy similar to single-factor experiments
- Can extend the range of validity and conclusions

Disadvantages of factorial experiments

- Indiscriminate use can lead to increased size, complexity and cost
- Too many treatments results in less precise estimates
- Factorial experiments have larger standard errors than single-factor designs

Types of factorial experiments

- Experiments with more than two factors
- Experiments with all factors at two levels – 2^k design
- Two-level factorials with blocking
- Two-level factorials without replication
- Fractional factorial experiments

Factorial example using two factors

Example:
What is the effect of yeast and potato starch on the loaf volume of sorghum bread?

Yeast levels of 2, 4 and 6% and potato starch levels of 10, 20 and 30% were used to make three loaves for each combination.

The loaf volumes were measured.

Combinations		Yeast		
		2%	4%	6%
Potato Starch	10%	A,B,C	D,E,F	G,H,I
	20%	J,K,L	M,N,O	P,Q,R
	30%	S,T,U	V,Z,X	Y,Z, A1

Resulting Loaf Volumes:

Combinations		Yeast		
		2%	4%	6%
Potato Starch	10%	450, 445, 452 (Mean=449)	495, 490, 485 (Mean=490)	500, 500, 489 (Mean=496)
	20%	525, 515, 510 (Mean=517)	500, 458, 499 (Mean=486)	595, 595, 594 (Mean=595)
	30%	595, 593, 592 (Mean=593)	593, 590, 589 (Mean=591)	570, 573, 574 (Mean=572)

Main effects

We want to know the ***main effect*** of yeast and the ***main effect of*** potato starch on loaf volume.

The ***main effects*** are the differences between the averages for each treatment level to the overall average.

(by how much does the loaf volume changes when using a specific treatment)

How ***significant*** are these effects?

Interaction

Interaction is the combined effect of two factors

We determine the average of one factor at each of the levels of other factors – if the increase (decrease) in average is similar between all subsequent levels, then there is no interaction

In our example we are interested in the interaction between yeast and potato starch and the effect of this interaction on loaf volume.

Also, how significant is this interaction?

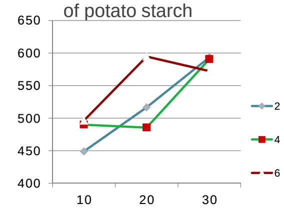
Cape Peninsula University of Technology
Centre for Postgraduate Studies

37/43

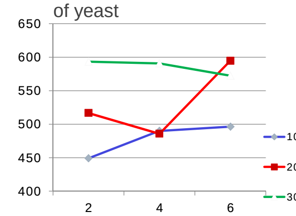
Interaction

Here the averages of one factor at each of the levels of the other factor are plotted

Average of yeast at each level of potato starch



Average of potato starch at each level of yeast



If there was no interaction between yeast and potato starch then the three lines in each graph would have been parallel.

Cape Peninsula University of Technology
Centre for Postgraduate Studies

38/43

First 2^k Factorial Design Example

A laboratory de-hulling machine fitted with eight abrasive discs, each 25 cm in diameter, driven by a variable speed electric motor was to be evaluated for performance in de-hulling cowpeas. A full factorial experimental design is to be adopted to evaluate the percentage yield with three variables at two levels:

Variable	Lower level	Upper level
Time (min) X_1	2	10
Batch size (kg) X_2	1.5	2.5
Speed (rpm) X_3	600	1400

Cape Peninsula University of Technology
Centre for Postgraduate Studies

39/43

Second 2^k Factorial Design Example

Table below presents the results of a 2^2 factorial design batter coating process with n = 4 replicates using the factors A = deposition time and B = flow rate. The two levels of deposition time are - = short and + = long, and the two levels of flow rate are - = 55% and + = 59%. The response variable is thickness (μm).

Treatment combination	A	B	Thickness (μm)
(1)	-1	-1	14.037, 14.165, 13.972, 13.907
A	+1	-1	14.821, 14.757, 14.843, 14.878
B	-1	+1	13.880, 13.860, 14.032, 13.914
ab	+1	+1	14.888, 14.921, 14.415, 14.932

Cape Peninsula University of Technology
Centre for Postgraduate Studies

40/43

Fractional Factorial Example

Five factors are suspected to affect the viscosity of a cereal beverage. Due to constraints in equipment and cost, the researcher needs to eliminate two of the five factors for inclusion in the final product formulation. The design matrix and the average viscosity score for each formulation (design point) is shown below. Which two factors will the researcher drop?

Design point	Factor					Response
	A	B	C	D	E	
1	+	+	+	-	+	7.2
2	-	+	+	+	-	6.5
3	-	-	+	+	+	4.0
4	+	-	-	+	+	3.8
5	-	+	-	-	+	4.1
6	+	-	+	-	-	3.4
7	+	+	-	+	-	6.0
8	-	-	-	-	-	3.0

http://www.phdcomics.com/comics/archive.php?comicid=413

References

- Weinberg, S. L. & Abramowitz, S. K. 2008. *Statistics using SPSS: An integrative approach,* New York, Cambridge University Press.
- Steffens, F. E. 1981. *Statistics STA204 (Design of experiments),* Pretoria, University of South Africa
- Hoshmand, A. R. 2006. *Design of experiments for Agriculture and the Natural Sciences,* Boca Raton, Florida, Chapman & Hall / CRC.